Information, Communication & Society

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rics20

Gov-tech as capture: public infrastructures under data capitalism

Burcu Baykurt

To cite this article: Burcu Baykurt (19 Mar 2025): Gov-tech as capture: public infrastructures under data capitalism, Information, Communication & Society, DOI: 10.1080/1369118X.2025.2479788

To link to this article: https://doi.org/10.1080/1369118X.2025.2479788

Published online: 19 Mar 2025.

INFORMATION, COMMUNICATION & SOCIETY
https://doi.org/10.1080/1369118X.2025.2479788

Routledge
Taylor & Francis Group

Gov-tech as capture: public infrastructures under data capitalism

Burcu Baykurt

Department of Communication, University of Massachusetts Amherst, Amherst, MA, USA

ABSTRACT

This paper examines the U.S. 'gov-tech' market, focusing on firms that partner with government agencies to redesign digital public infrastructures. Through interviews with founders and chief technology officers of gov-tech companies, it develops 'capture' as a framework for analyzing how these firms commodify public data, redefine state capacity, and foster dependencies through computational systems. The study identifies three mechanisms of capture: grabbing and aggregating public datasets (data capture), reconfiguring public services into computable units (value capture), and transforming public data into market assets (regulatory capture). Drawing on insights from critical data studies and political economy, the paper demonstrates how gov-tech firms redefine state capacity and public accountability while examining the broader implications of embedding private interests in public governance under data capitalism.

ARTICLE HISTORY
Received 13 July 2024
Accepted 10 March 2025

KEYWORDS
Government; algorithms; data assets; capture; surveillance

Introduction

'I assume we are all in this room because we want to answer the question of how we become a data-driven city?' began Lucas, scanning the room. We were at a daylong workshop where public officials from several cities gathered to talk about smart cities and data-driven policymaking.[1] Lucas was the CEO of a data analytics company in the United States (U.S.), one that worked with local governments on data-centric projects such as dashboards, smart city pilots, or predictive analytics programs. After letting his initial question settle, he added, with deliberate emphasis, 'It *is* work!' Heads in the room nodded in agreement. Lucas went on to divide data-driven cities into two groups: ones having 'actual pain points that need to be addressed' and those that launch smart city initiatives or other programs and hope that 'somebody is going to come and figure out what to do with this data.' Laughter rippled across the room. Playfully, he added, 'We spend a lot of time and beer to help you figure out what to do with data.'[2]

Lucas's company, a relatively small startup in the U.S. Midwest, is an example of a gov-tech ('government technologies') firm. It works with municipal or state agencies, designing dashboards for them, digitizing existing maps or other paperwork, and advising public organizations about how to work with (big) data and machine learning. 'Gov-tech' is a term often used to describe various technologies adopted within public agencies

CONTACT Burcu Baykurt bbaykurt@umass.edu

(Gordon, 2021) but it increasingly refers to a sub-set of tech companies that provide government agencies with expertise in data science and software for analyzing and visualizing public datasets (Bharosa, 2022). Some of these firms are niche and tackle one specific issue, such as curb management or property code enforcement. Others offer a broader range of services, addressing issues such as housing, transportation, or public health. Gov-tech firms peddle their services as a means to streamline data-driven decision-making and build civic information tools for the public. As more government agencies try to use automated, predictive, or simply data-driven decision-making systems, they turn to gov-tech firms.

This paper examines gov-tech firms as a case in how state capacity and tech capital intertwine in new ways under data capitalism. Based on semi-structured interviews conducted with founders, chief executive officers, and chief technology officers of gov-tech startups in the U.S., it investigates how these companies find ways to extract value from their growing role in state administration. The argument hinges on the concept of *capture*, in contrast to surveillance, in capitalism. Gov-tech firms, the paper argues, are less interested in expanding the state's monitoring capabilities than in reshaping existing public services and information so that they can be incorporated into – and be governed by – computational systems. This reframing shifts the conversation around the tech industry's relationship with government from a focus on spectacle or intrusion to something more mundane but equally transformative.

The paper, therefore, offers two contributions to the fields of political economy of digital technologies and critical data studies. First, it positions capture as a practical, everyday framework for understanding how digitalization turns public administration into a new market frontier, one that generates profit by embedding tech firms into the state's infrastructure. Second, it argues that this process gives regulatory capture an updated meaning, as gov-tech firms position themselves as intermediaries between public agencies and different kinds of markets, profiting from the monetization of public datasets.

The rest of the paper first situates gov-tech firms in the overall political economy of the datafied state and then introduces the theoretical purchase of capture for understanding the practices of this particular subset of companies. Following a brief discussion of methods, it lays out three mechanisms through which capture unfolds: (1) grabbing and digitizing existing public datasets, (2) reframing public services and information into new categories, and (3) stretching the practices of regulatory capture into new forms of influence. The discussion suggests that capture as a framework identifies the productive (in a capitalistic sense) link between trapping or preserving public datasets and transforming them into new economic and political values. It concludes by raising questions about what becomes of public information as it is recast from digital public goods in the datafied state into data assets.

The political economy of the datafied state

The 'datafied' state can be defined broadly as the centering of data in governance wherein state agencies monitor, evaluate, and control citizens and territories via data and databases (Cheung & Chen, 2022). Unlike a simple conversion of paper records to digital formats (Datta, 2023), however, the datafied state indicates a more significant change in how government agencies operate and make decisions (Fourcade & Gordon, 2020; Johns,

2019). This shift, preceded by the neoliberal transformation of the state, is characterized by the introduction of a new group of actors, such as tech companies and civic technologists, and using new techniques, such as machine learning, in governance. These changes vary across a spectrum of democratic and authoritarian contexts as well as various levels of government offices, often contributing to the reduction of government resources and the obfuscation of a given government's ties with and responsibilities toward markets.

Studies of the datafied state examine the impact of data-centric technologies and algorithmic techniques on enhancing the state's legibility and its capacity to scale up and centralize decision-making (Burrell et al., 2024). Critical analyses underscore concerns regarding privacy invasion, bias amplification, and erosion of accountability in policy-making (Brayne, 2020; Eubanks, 2018; Pasquale, 2016). Others point to deeper shifts in how the state sees, identifies, and learns public concerns (Crooks, 2022; Fourcade & Gordon, 2020; Johns, 2019). The new techniques and tools of the datafied state, such as dashboards or predictive analytics, transform the day-to-day operations of government agencies. They also fundamentally redefine some of the public responsibilities of the state by turning them into computing tasks. Social welfare, for instance, becomes a task of optimizing public funds rather than a redistribution policy and delivery of social services (Dencik & Kaun, 2020).

In addition to the profound changes in citizen-state relationships, in which citizens have limited autonomy against the expansion of the datafied state, the state's capacity to serve the public interest is also diminished. A new group of 'para-state' contractors increasingly take on public roles and responsibilities (Datta, 2023; Dunleavy et al., 2006; Dunleavy & Margetts, 2024; Taylor, 2021). From tech giants in Silicon Valley to local and transnational tech firms providing digital infrastructures, public services are contracted out to tech companies, administrative processes are automated, and governments enter into data and revenue-sharing agreements with tech firms (Lopez-Solano & Castañeda, 2024; Luque-Ayala & Neves Maia, 2019; Mahase, 2019). Increasingly, public services and government data become a site of value-creation for the tech industry (Reutter & Åm, 2024). Instead of regulating these markets, governments often facilitate their growth (Baykurt, 2024).

One of the dominant frameworks to explain this evolving relationship between the state and tech companies is 'surveillance capitalism' (Sadowski, 2020; West, 2019; Zuboff, 2019). Putting the emphasis on the intensification of surveillance technologies and behavioral manipulation, this framework highlights the ways people are monitored, classified, sorted, and ranked via digital technologies. By expanding the surveillance apparatus beyond the confines of the state, the new 'automated surveillance' – or 'big data surveillance' – controls populations and environments by indiscriminate monitoring and further obscuring the accountability of both governments and corporations (Andrejevic, 2019; Andrejevic & Gates, 2014; Murakami Wood & Steeves, 2021).

While this interdisciplinary scholarship on the datafied state and surveillance capitalism is illuminating, it also presents some limitations, particularly considering the expansion of the gov-tech market. Partnerships between governments and tech companies, such as smart cities (Barns, 2020) or gov-tech contracts, do not always align with the all-encompassing surveillance paradigm. Nor do they operate as seamlessly as social media platforms that strive to customize their services for individual users. Instead,

what often emerges is partial, unfinished, and uneven (Murakami Wood & Mackinnon, 2019). The fragmented structure of various government units and regulations, combined with the challenge of trying to achieve the multiple goals of public institutions, does not allow for the kind of continuous, 'frameless' monitoring associated with surveillance capitalism (Andrejevic, 2019). This evolving assemblage – or what Stephanie Sherman (2023) calls the 'polyopticon' – instead generates incomplete, diverse, and often ambiguous insights, which can sometimes be 'unintelligible to humans' (pp. 1217) How to then elucidate the logics and practices of the new partnerships between states and tech companies? The following section proposes one potential approach.

Toward a capture model of the datafied state

An alternative political-economic analytic to understand the evolving practices of public agencies trying to become data-driven is *capture*. In an influential essay published in 1994, Philip Agre distinguishes between surveillance and capture as two models of power in a digital society. Capture, a key practice in computing, refers to the (1) acquisition of data and (2) modeling or translation of the reality reflected by data (Agre, 1994). Unlike constant monitoring and recording of new information (surveillance), capture emphasizes the process of parsing human activities into computer languages and the subsequent interventions to facilitate it. Behind this seemingly technical practice is a set of social relations, narratives, and values that constitute the collective political work of reconstructing social life into units to be captured by computing.

Centering capture when examining the datafied state demonstrates how public administration can be broken into parts that can be handled by gov-tech firms. While still implicated in monitoring and recording, many of these companies focus on acquiring existing datasets in a public agency rather than generating new information. As Reutter and Åm (2024) observe in the case of Norway, in some cases, 'the sheer availability of quality data in a workable format' (pp. 15) is considered enough for tech companies and policymakers to expect innovation and revenue generation.

Gov-tech firms often propose to repurpose existing public data and services into new analytical products for the market, earning money from public funds and tying government agencies to their platforms through subscription fees. In the gov-tech market, the ability to turn public information into units that computers can work with – what Agre (1994, p. 110) called 'grammars of action' – is as important, if not more so, than simply tracking and recording information (Beverungen, 2021).

Using capture as a framework allows for understanding how gov-tech companies create value in this new market and raises new research questions. What kind of data are amenable to be captured and become valuable on the gov-tech market? How, if at all, does this imperative to capture change the way gov-tech firms and public agencies work together? Perhaps most importantly, what happens after data is captured? How do computation and the logics of data capitalism transform public information into new forms and values? While this paper does not offer exhaustive answers to these questions, it posits that the concept of capture provides a useful lens to understand the partnerships between gov-tech companies and public agencies, one that sheds light on the often-mundane ways government services and the interactions between citizens and the state are transformed and monetized by the tech industry.

Methods

This paper draws on twelve semi-structured interviews conducted with founders, chief executive officers (CEO), or chief technology officers (CTO) of gov-tech firms based in the U.S. The sample of tech start-ups specializing in data science and contracting with public agencies relies on several sources, including the LinkedIn networking platform, databases of early-stage startups, and publicly available 'Top 100' lists from the website govtech.com. The initial sample consisted of ninety-five gov-tech companies. In contrast to major tech corporations like Sidewalk Labs, IBM, or Tyler Technologies, this study focused on smaller to mid-sized startups in order to secure relatively easier access and to understand how the industry perceives where innovation lies. After excluding those that specialized solely in policing (due to concerns that they might be less willing to participate in interviews), those that provided only network security services, and startups that had become defunct by the time we initiated the interviews, we had a final list of twenty-five companies. At these gov-tech firms, we contacted their founders, CEOs, or chief data officers between September 2021 and March 2022 (Table 1). The remaining firms either did not respond to our requests or were unable to find the time to talk to us over a nine-month period.

The interview questions focused on four key areas: (1) the founding stories of these firms, including their motivations for entering the gov-tech market; (2) the products and services they offer to public agencies; (3) how they communicate the value of their business to public agencies as well as investors; and (4) their work processes with public agencies. Prior to each interview, we reviewed the websites and online demos of each firm to understand their products and their client base. We slightly revised interview questions to reflect each company's area of expertise. On average, these interviews lasted slightly over an hour and a half and were conducted over Zoom.

To analyze the interviews, we used open coding (Corbin & Strauss, 2015) to identify how these firms described the needs in the market that drove them to launch gov-tech firms, and the ways they built partnerships with public agencies and set up work routines. We noted the similarities and differences across interview transcripts to categorize the recurring themes. As initial themes emerged, we developed an empirically grounded theory (Charmaz, 2006) of capture as the organizing principle of the gov-tech market to elucidate the political economy of gov-tech firms and their partnerships with public agencies. The paper mentions company numbers and the role of the interviewees in each firm to help the reader connect quotes from the same person.

Table 1. The list of interviewees and dates of interviews.

Company #1	CTO	September 2, 2021
Company #2	Co-Founder	September 13, 2021
Company #3	CEO	September 9, 2021
Company #4	CEO & Founder	March 23, 2022
Company #5	CEO & Co-Founder	September 22, 2021
Company #6	CEO & Co-Founder	March 24, 2022
Company #7	CEO & Co-Founder	April 26, 2022
Company #8	CEO	May 3, 2022
Company #9	CEO & Co-Founder	April 22, 2022
Company #10	CEO	April 21, 2022
Company #11	Founder	April 19, 2022
Company #12	CEO & Co-Founder	April 25, 2022

Gov-tech: capture in three acts

This section traces how capture unfolds in the relationship between gov-tech firms and public agencies. It begins with data-grabbing, that is, acquiring publicly available datasets. From there, it moves onto reorganizing government records, shaping them into new 'grammars of action,' discrete units that can be repackaged and resold. What once belonged to the public becomes a proprietary asset, shared with other public agencies and market actors as a new source of information. In the final stage, gov-tech firms, having embedded themselves in the infrastructure of governance, desire to leverage the captured data not just for profit, but for power – turning data into a form of political currency, one that allows them to influence other industries, shape policy, and ultimately dictate the terms of the systems they claim to serve.

Grabbing mundane data

Gov-tech firms rely on two crucial trends that have shaped the datafied state over the last few decades: the platformization and datafication of government services (Burrell & Singh, 2024). The modern state has always been a calculating machine. But from the much-hyped smart cities to the platformization of government services, public agencies have further solidified the technical infrastructure of the state. The business strategy of gov-tech firms plays to this deepening reconstruction of the state through numbers and, in particular, targets big tech companies' recent efforts to implement datafied systems in government.

Tech giants have a reputation for treading heavily and pitching imperfect, glitch-ridden software projects. Small-to-mid sized gov-tech companies, however, present themselves as a more responsive alternative path to datafied governance. In interviews, for example, gov-tech founders were quick to draw a line between their firms and Big Tech. Many boasted of past lives in civic tech or stints in local government, such as urban planning, education, or public health. At a minimum, each could point to a story around a local inefficiency that sparked their entrepreneurial interest. They were not the typical Silicon Valley types, they insisted, parachuting in with lofty promises. They were motivated by a desire to work within the system. As one founder put it, their company's mission was simple: ensuring 'people who work in city government to have the tools they need to do their job effectively.' He then continued:

> I don't think having more cameras in the world will make government more effective. I don't think that, you know, the blockchain will make the government better at fixing the stairs in the park. But I do think that having government officials who can do their jobs for the government would be fantastic. And I think giving them better software tools can be a big part of that (Company 1, CTO).

Another co-founder of a gov-tech startup contrasted their work with smart cities:

> [Smart cities] are companies selling dashboards of masses, everything is connected, and all the data is flowing perfectly. What we represent is a much more organic approach to modernizing cities, at a per service level. As individual services will improve interaction, day to day, with your city, that rebuilds confidence in government (Company 2, Co-Founder).

Gov-tech firms position themselves as an intermediary between Big Tech and public institutions. Their focus, they assert, lies not in the industry's glitzy visions of 'smart

cities' (a term studiously avoided for its association with surveillance and past failures) but in the unglamorous realities of public service. User-friendly design, data science expertise, and a commitment to civic values take center stage here. These firms do not peddle the hype of future innovation (Russell & Vinsel, 2018) but claim to tackle the problems of present inefficiencies and serve the public good while arming public officials with what they believe to be a messy but essential tool: better and bigger data.

The gov-tech playbook, therefore, always opens with a capture of mundane data. Once they begin to work with a public agency, gov-tech firms aim to grab all public datasets, even those that have yet to be digitized such as old maps or regulatory documents. The objective is to build an inventory of all the kinds of datasets (or 'data assets,' as many call them) the agency has. From websites to demos to sales pitches, the gov-tech mantra is to track existing data, digitize paperwork, maintain data operations, and troubleshoot. This initial act of capture – the data grab, the meticulous blending, the user-friendly interface – was a constant across all the firms interviewed. A gov-tech CEO insisted that '60% of [our] work is implementation and 40% is just maintaining data pipelines' (Company 5, CEO & Co-Founder).

In interviews, gov-tech firms were careful to tie the need for data-grab to the public interest and equity objectives of government agencies. One gov-tech CEO, for instance, stressed that 'Wanting to overlay all [data] with socio-economic and demographic information is becoming more important, as more cities recognize, as more cities sort of apply an equity lens to what they're doing' (Company 5, CEO & Co-Founder).

Another gov-tech CEO drew a connection between the imperative for local governments to use data analytics and their desire to advance social justice in cities:

> Our data is really well suited to help [public agencies] deal with systemic problems because they have a lot to do with the conditions of the communities you are working in. [Public officials] need to layer environmental data and socioeconomic data; that's actually kind of hard to do. But that's something we do really well. If I want to grab a fifty-year flood plan and use the environmental justice index or the census tracts that are eligible for low-income housing tax credits, I can do that. [For any project], public officials need to answer: Are you implementing that equitably? How are you communicating the return on investment? Do your electeds know the impact of it, and relatively to these other things? (Company 3, CEO).

Beyond the allure of taking over tedious data work for resource-poor public institutions, gov-tech firms lay claim to another skill: data wrangling. They offer to track down public and private datasets, scrape them, and present them to public officials, who ostensibly may not even be aware of the data's existence, let alone how to use it to make decisions. 'We offer a platform with a common data model,' explained one CEO (Company 6, CEO & Co-Founder), 'allowing you to ingest multiple data points.' One founder suggested that most public agencies possess ample data, but they do not know how to use it:

> I have a lot of contacts in the mayor's office in San Francisco and San Jose. In talking to them, I realize that they have a truckload of data, but they don't have the right data. And they don't know what to do with the data they have. And they don't know how to procure the right data. It was a complete mess (Company 4, CEO & Founder).

In the Silicon Valley narrative, public institutions are cast as bumbling bureaucracies incapable of fulfilling their mandates, a convenient backdrop against which technology

swoops in as a neutral arbiter. Gov-tech firms, however, chart a different course. They eschew a disruptive role, opting instead to focus on the routine drudgery of government data work. Contrary to the usual dichotomy in critical technology studies between innovation and maintenance (Mattern, 2018), gov-tech firms claim to rally data in the service of repair, maintenance, and management of the creaking infrastructure of public data. Their promises are grounded in the digitization, integration, and migration of existing datasets.

By emphasizing and taking on this often laborious task of grabbing, cleaning, reorganizing, and collating information, gov-tech firms believe they form alliances with public agencies and articulate a much more palatable value proposition. Their work recognizes government datasets as digital public goods and ostensibly aims to improve them. But capture is not a passive act. It also kicks off the process of transforming public information into a potentially valuable asset in the data economy.

Bundling disparate data

If the first step of capture is the data grab, a review and collection of information that public agencies already have, the second step is reshaping this information to fit the needs of a government agency. Here, the data sets are blended or broken into new units, in the gov-tech speak, to become valuable tools for decision-making. Census figures and a city's housing inventory, for instance, are mixed to inform policies on affordable housing. This granularization of existing information and refinement of categories, achieved by weaving in other public datasets, demonstrates how gov-tech firms not just grab data but also transform them into new units to extract value from and facilitate further capture of government datasets.

One company CEO described their platform as a central repository of a public agency's data and a gateway to access diverse external datasets. 'Maybe it is socioeconomics, demographics, crashes, foot traffic, or congestion,' he said, 'we have dozens of different data sources that are available [for any county or city]' (Company 6, CEO & Co-Founder). Their pitch to clients was to either 'grab that data' or have the company 'deliver it inside [their] data.' A third option was to do more analysis, such as performance management, and even do it in real time. Interviews with gov-tech firms revealed that they present this data manipulation as a value proposition, repackaging extant public information for a fee. Often, this involves tiered subscription models, where flashier datasets may come at a premium.

An exchange on a gov-tech firm's LinkedIn page illustrates this dynamic clearly. The company, at the time, was announcing the launch of county-level health rankings in the U.S. with the additional note that it has the 'country's largest library of neighborhood data that is complete, consistent, and connected.' One user, a self-identified 'applied sociologist,' inquired if the company's data 'was free to download or is there a charge?' 'County health rankings are free,' this user's comment noted, 'And of course, data.census.gov is free.' The response, delivered with the promotional tone of a corporate account, touted the company's 'pretty special' data library, 'designed for people who need something useful very quickly, at granular levels, including custom geographies.' And, predictably, it ended with an invitation to check out the company website. Then the company CEO chimed in, 'One of my favorite things about the work we do,' he

said, 'is when we take an important public dataset and give it superpowers!' In his response, the original commenter was persistent: 'But you didn't answer my question. Is it free to download, or is there a charge?'

The answer, of course, was no; the product was not free to download. This brief exchange lays bare a core tenet of gov-tech firms: the art of re-packaging. Their service depends on the act of aggregating disparate datasets, dressing them in a visually appealing interface, and, most importantly, maintaining – for a fee, of course – this curated information repository. The silver bullet is not some predictive analytics but this banal act of tailoring existing public data, like the census or Bureau of Labor Statistics reports, to specific locales and agencies. As one gov-tech CEO candidly explained, what they do 'doesn't have to be impressive and sophisticated' (Company 3, CEO). He then offered a telling example: a city reaching out to the firm to seek solutions for declining population and affordable housing. 'There's this employment dataset, and we have the projections on the number of housing units you have and how many of them are affordable,' he explained, 'So this is actually a simple math equation that you're not doing right now. Because you don't know which data sets are out there.'

The implication here is that public institutions lack the wherewithal to navigate the complexities of working with public datasets. But gov-tech firms exploit more than just a lack of technical expertise. They position themselves as mediators navigating the bureaucratic chasms that separate public institutions. 'Consider the divide between the court system and a municipal government,' one CEO offered by way of illustration. 'In a city that deals with an epidemic of evictions,' he explained, 'you would think that housing courts would share that data with the mayor of the city' (Company 5, CEO & Co-Founder). He then elaborated:

> But the data is out there, and we can scrape it because it's all a matter of public record. We can do a quick search and see their entire court docket. But they are not going to give bulk data like this to local governments. Sometimes they don't want to share data, but we've got some pretty crafty folks on our team who will figure out … So, we are playing that end-to-end function. We build the tools into our front end to make it very easy to find [data]. And on the back-end, we are doing this frenzied work of data aggregation and cleaning (Company 5, CEO & Co-Founder).

Gov-tech firms thus thrive on these bureaucratic shortcomings – the fragmented systems, the petty rivalries, and the budget constraints that prevent public agencies from hiring the talent needed to wrangle data into usable form. These services of chasing, grabbing, and blending data are offered to government agencies via a subscription model with different datasets triggering different price points. As one gov-tech founder explained, 'If a city is willing to pay for property sales data, we flip the switch and turn that feature on quickly. If you can't get pre-foreclosure data, we can do the same. If you want vacancy data from the US Postal Service, we can turn that on and bring that value-added data to our local governments' (Company 5, CEO & Co-Founder). Gov-tech firms are adept at navigating these information asymmetries, carving out a niche in the midst of jurisdictional quagmires.

While each gov-tech firm describes what they do as a form of expertise in interpreting numbers, they are also careful not to frame themselves as consultants. As one founder explained, 'Government is so used to hiring professional services firms or engineering

firms who act as consultants. But what they give them is a stale, static PDF, which often dies' (Company 6, CEO & Co-Founder). What gov-tech firms offer, in this founder's description, is 'the same proficiency, but [delivered] in a scalable platform.' In other words, the process of collaborating and what they consider to be 'generating value' comes back to the same model: 'combine the datasets, clean it a bit, structure it, apply the machine learning algorithms, and then produce the insights that go into the dashboard' (Company 1, CTO). Gov-tech, as discussed in the previous section, is not about groundbreaking innovation; it is about monetizing the repackaging of public information.

Operating as part consultant, part data scientist, and part digital detective, gov-tech firms marshal capture both in terms of grabbing data and in blending and translating disparate datasets into new categories and metrics. These two steps ultimately aspire towards a third form of capture: not just of data, but of the very way public agencies conceive of governing. The following section focuses on this third form, a kind of capture similar to corporate capture of the datafied state but executed more creatively than traditional regulatory capture.

Corporate capture

Just like every start-up, gov-tech firms keep their eye on scaling up. But how? One of the gov-tech firms explained that their challenge in the larger start-up ecosystem is 'to accelerate the growth while working with governments at their pace,' and then he added, 'There is also a fairly finite list of cities or agencies' (Company 8, CEO). Still, gov-tech firms mostly bet on the 'enormous value of aggregated data' (Company 4, CEO & Founder), culled from a multitude of sources and served up on their curated platforms. The dream: becoming the 'Google of [X],' be it transportation, housing, or water data, a one-stop shop for government data in a specific domain.

One co-founder talked about venturing into 'indexes or predictions' to illuminate larger trends within their domain (Company 2, Co-Founder). Another founder and CEO envisioned a platform with different applications from the same datasets:

> We are trying to become the source of truth and a platform that offers multiple use cases or applications from the same data. You can take our dataset and start your mobility project. Or you can take the dataset and start an economic development activity, or disaster or resilience initiative. Whatever it is, you need those core source datasets, so we are trying to be that source across many regions (Company 4, CEO & Founder).

Other gov-tech firms were more specific about the way they wanted to present the value of the aggregated data captured from public datasets: 'Once we are managing data for lots of different cities, we can be the data platform that the private sector uses to access this data' (Company 1, CTO), one chief technology officer remarked. This founder elaborated on their vision to play this mediator role between the public and the private sector:

> Every city wants people using their transportation systems to understand it better. And particularly, they want people to respond to their regulations. That's a very nice sales pitch both for the private sector and for the city (Company 1, CTO).

Real estate firms, some gov-tech firms suggested, if their data leaned towards property, would find their platform invaluable. Or fleet companies could become prime clientele if

the firm's focus was on transportation or curb management. Sociologists Marion Fourcade and Jeff Gordon (2020) describe this process as 'private appropriation of public data' wherein tech firms extract from state-collected or owned information. But the gov-tech vision is not a simple case of privatization; it is also developing a new technique of building influence about and toward government agencies, one that could be valuable for market actors. Gov-tech firms aim to use the captured data to mediate between markets and governments. By controlling access to knowledge about specific industries, gov-tech firms desire to shape data flows and manage the relationship between public and private actors.

Given the size and capacity of gov-tech start-ups, this third kind of capture may seem a distant prospect. At the federal level, however, one illustrative case offers a glimpse into this potential future. The company CostQuest describes itself as a 'broadband consulting firm … developing GIS data for the broadband ecosystem.' In 2022, the Federal Communication Commission (FCC) entrusted CostQuest with the critical task of revamping national broadband maps, which are foundational tools for allocating federal broadband deployment funds. This map relied on data received from the FCC itself, along with various public datasets like land records and address databases culled from public agencies across the country. CostQuest asserted that it combined these with 'internal data and machine learning models' (Engebretson, 2023) to create the final product, the proprietary FCC National Broadband Map.

The map created a public controversy as both public and private institutions required a license for access (Ali, Forthcoming; McGarry, 2022). The FCC later tried to mitigate this crisis by offering free access in different use cases, but the fundamental conflict remained intact: public agencies' data created an information commodity to be sold to public agencies and the industry actors they tried to regulate.

This ambition of gov-tech – capturing the datafied state – is similar to regulatory capture, in which private companies influence the rules and limits of regulation, not just through lobbying or donations, but by shaping how regulations are understood and applied (Carpenter & Moss, 2013). Gov-tech firms, by leveraging publicly available data and their expertise in data science, aspire to develop epistemic capacity and play this intermediary role between markets and governments. They aim to shape how government data are used and interpreted within specific industries. By controlling access to and shaping the interpretation of public data, gov-tech firms could, in the long run, impact how public information is collected and curated inside government agencies in ways that could benefit their own growth and client base. This would ultimately steer public discourse and decision-making, aligning them more likely with market interests rather than the public good.

Discussion

In critical studies of technology, the term capture has come up recently to describe the tech industry's capture of the state (Whittaker, 2021) or to discuss the rise of 'captology' in algorithmic recommender systems (Seaver, 2019). The concept reflects a growing focus on how tech companies shape the systems they operate within, driven by their own business interests. Building on these critical interventions, this paper uses the case of gov-tech companies to further theorize the mechanisms of capture in mediating

the relationship between the state and markets. Gov-tech firms demonstrate that capture in data capitalism is not only a means of trapping or preserving data but also of taking and transforming them.

Examining capture as a defining element of capitalism, Carolyn Hardin (2021), drawing on Deleuze and Guattari (2007), argues that capture creates the possibility of comparison that constitutes rather than discovers surplus, thereby enabling its monopolistic appropriation. By taking various parts of public data infrastructures – open datasets, government record logs, digitized maps, and regulations – and turning them into computational products, gov-tech firms produce new conditions for comparison across various kinds of information and attempt to generate new value. Both technically and metaphorically, government datasets morph into traces to be captured by gov-tech firms, fragmented into new units, and bundled and re-bundled based on the expectations of public and private clients (Birch & Muniesa, 2020). Gov-tech firms aspire to configure the curated platforms of government data into novel resources, over which they claim ownership and extract revenue through licensing. In this process, the very definition of public information and the meaning of regulatory capture undergo a radical shift.

Gov-tech firms, in constructing the new architecture of public data, do more than forge dependencies for public agencies (Ashuri, 2023); they also aspire to reconfigure existing datasets into categories and metrics that hold value for a range of industries. By blending the computational capture and value capture with regulatory capture, gov-tech companies aim to make new markets *and* expand the political meaning of capture. This expansive version of capture and the concomitant transformation of public data infrastructures into a trove of data assets is not a simple case of privatization of public goods. It compels us to ask what happens to public services as they are reshaped for extraction and to align with the broader commercial objectives of gov-tech firms that cater to private markets.

This paper exclusively focuses on small and medium-sized tech companies. Though they are small, their endgame is often an acquisition by one of the larger tech companies in Silicon Valley. In other words, they depend, in ways both subtle and obvious, on the ecosystems controlled by Big Tech – cloud computing platforms, data analytics tools, and the infrastructural scaffolding that underpins their operations (Langley & Leyshon, 2020; Narayan, 2022). The political and economic power of Big Tech is poised to shape both the development and the limitations of these mechanisms of capture within the gov-tech market. Gov-tech firms may currently seem small, even scrappy, compared to Silicon Valley giants, but their practices foreshadow how the state itself becomes a source of capital.

Conclusion

The rise of gov-tech start-ups in the United States has been quiet, almost imperceptible – an industry made up of small to mid-sized firms that present themselves as far removed from the innovation culture of Silicon Valley. Their mission, as described by founders and CEOs, is more modest, more practical: to aid resource-poor public agencies by organizing data infrastructures and providing them with a wider range of information for decision-making. These companies promise to solve the problems of government

bureaucracy – not through disruption, but through quiet collaboration. Yet beneath this rhetoric of utility is a deeper mechanism of value creation, at the heart of which lies *capture*.

Capture, I have argued, refers to a series of moves that define the gov-tech business model. First, these firms take publicly available datasets – often incomplete, unstructured, or dormant – and aggregate them. Next, they transform these datasets into proprietary products, packaged and licensed for government use. Finally, they leverage this knowledge to expand into new industries and develop a new kind of political influence. Gov-tech firms, in this sense, are not just service providers working at the periphery of public agencies. They play a central role in restructuring the government's own data – taking what is diffuse, fragmented, or free, and rendering them newly valuable in the market.

Despite the relatively small sample size, the concept of capture and its mechanisms were consistently observed across the gov-tech firms interviewed in this study. Expanding the pool of firms would help uncover variations based on company size, areas of expertise, and types of clients. Plus, given the significant political and economic shifts in the United States under the second Trump administration – and the ongoing evolution of the tech industry's capture of the datafied state – more research is needed to identify both continuities and emerging differences in the larger gov-tech market.

Future research should also explore the ramifications of capturing public information. How do public officials perceive the role of data in decision-making when it is reconfigured for profit-seeking by gov-tech firms? What ethical norms and obligations arise as these firms seize and re-value government datasets? To what extent is the rise of the gov-tech market and the transformation of public data into assets a U.S. phenomenon? If the gov-tech market expands to other countries, what would be the similarities and differences? Finally, more research is necessary to identify the specific ties between this bourgeoning market and Big Tech, and the ways these bigger players shape the different mechanisms of capture. Answering some of these questions offers a better understanding of the relationship between public information and the tech industry's shifting discourses of big(ger) data.

Notes

1. All names of people and companies are anonymized.
2. The event took place in a Midwestern city in the U.S. on October 26, 2016.

Acknowledgments

I extend my gratitude to the Edgelands Institute, the Weizenbaum Institute, and the Global Humanities Institute 2024: Design Justice AI for funding, invaluable camaraderie, and the intellectual space they provided to complete this research. I am especially thankful to Kevin Zheng, Annika Our, and Robert Long for their outstanding research support. Lastly, this work benefited from discussions at the Oregon Surveillance Studies Workshop in 2023 and the EASST-4S conference in 2024.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Burcu Baykurt is an assistant professor of media studies at the University of Massachusetts Amherst. Her research examines how digital infrastructures reshape and perpetuate durable inequalities. [email: bbaykurt@umass.edu].

References

Agre, P. E. (1994). Surveillance and capture: Two models of privacy. *The Information Society*, *10*(2), 101–127. https://doi.org/10.1080/01972243.1994.9960162

Ali, C. (Forthcoming). Where the market dares not tread: Mapping rural broadband in the United States. In P. Brodie & D. Barney (Eds.), *Media rurality* [Manuscript submitted for publication]. Duke University Press.

Andrejevic, M. (2019). Automating surveillance. *Surveillance & Society*, *17*(1/2), 7–13. https://doi.org/10.24908/ss.v17i1/2.12930

Andrejevic, M., & Gates, K. (Eds.). (2014). Big data surveillance: Introduction. *Surveillance & Society*, *12*(2), 185–196. https://doi.org/10.24908/ss.v12i2.5242

Ashuri, T. (2023). Shadowy knowledge infrastructures. *Information, Communication, & Society*, *27*(3), 583–599. https://doi.org/10.1080/1369118X.2023.2227667

Barns, S. (2020). *Platform urbanism: Negotiating platform ecosystems in connected cities*. Springer Nature.

Baykurt, B. (2024). Corporate capture. In J. Burrell, R. Singh, & P. Davison (Eds.), *Keywords of the datafied state* (pp. 49–58). Data & Society Research Institute.

Beverungen, A. (2021). The invisualities of capture in Amazon's logistical operations. *Digital Culture & Society*, *7*(2), 185–202. https://doi.org/10.14361/dcs-2021-070209

Bharosa, N. (2022). The rise of GovTech: Trojan horse or blessing in disguise? A research agenda. *Government Information Quarterly*, *39*(3), 101692. https://doi.org/10.1016/j.giq.2022.101692

Birch, K., & Muniesa, F. (Eds.). (2020). *Assetization: Turning things into assets in technoscientific capitalism*. The MIT Press.

Brayne, S. (2020). *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press.

Burrell, J., & Singh, R. (2024). Introduction: Reimagining the state in a data-driven world. In J. Burrell, R. Singh, & P. Davison (Eds.), *Keywords of the datafied state* (pp. 19–32). Data & Society Research Institute.

Burrell, J., Singh, R., & Davison, P. (Eds.). (2024). *Keywords of the datafied state*. Data & Society Research Institute.

Carpenter, D., & Moss, D. A. (2013). *Preventing regulatory capture: Special interest influence and how to limit it*. Cambridge University Press.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage Publications.

Cheung, A. S. Y., & Chen, Y. (2022). From datafication to data state: Making sense of China's social credit system and its implications. *Law & Social Inquiry*, *47*(4), 1137–1171. https://doi.org/10.1017/lsi.2021.56

Corbin, J., & Strauss, A. L. (2015). *Basics of qualitative research*. SAGE Publications, Inc.

Crooks, R. (2022). Seeking liberation: Surveillance, datafication, and race. *Surveillance & Society*, *20*(4), 413–419. https://doi.org/10.24908/ss.v20i4.15983

Datta, A. (2023). The digitalising state: Governing digitalisation-as-urbanization in the global south. *Progress in Human Geography*, *47*(1), 141–159. https://doi.org/10.1177/03091325221141798

Deleuze, G., & Guattari, F. (2007). *A thousand Plateaus: Capitalism and Schizophrenia*. Translated by Brian Massumi. Minneapolis: University of Minnesota Press.

Dencik, L., & Kaun, A. (2020). Datafication and the welfare state. *Global Perspectives*, *1*(1), 12912. https://doi.org/10.1525/gp.2020.12912.

Dunleavy, P., & Margetts, H. (2024). The political economy of digital government: How silicon valley firms drove conversion to data science and artificial intelligence in public management. *Public Money & Management*, Advance online publication, http://doi.org/10.13140/RG.2.2.33965.70887

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New public management is dead – long live digital-era governance. *Journal of Public Administration Research and Theory*, *16*(3), 467–494. https://doi.org/10.1093/jopart/mui057

Engebretson, J. (2023, March 3). *EXCLUSIVE: CostQuest breaks its silence on broadband map issues*. Telecompetitor. https://www.telecompetitor.com/exclusive-costquest-breaks-its-silence-on-broadband-map-issues/.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Fourcade, M., & Gordon, J. (2020). Learning like a state: Statecraft in the digital age. *Journal of Law and Political Economy*, *1*(1), 78–108. https://doi.org/10.5070/LP61150258.

Gordon, E. (2021). Civic tech/gov tech. In P. Harris, A. Bitonti, C. S. Fleisher, & A. S. Binderkrantz (Eds.), *The palgrave encyclopedia of interest groups, lobbying and public affairs* (pp. 1–4). Springer International Publishing.

Hardin, C. (2021). *Capturing finance: Arbitrage and social domination*. Duke University Press.

Johns, F. (2019). From planning to prototypes: New ways of seeing like a state. *The Modern Law Review*, *82*(5), 833–863. https://doi.org/10.1111/1468-2230.12442

Langley, P., & Leyshon, A. (2020). The platform political economy of FinTech: Reintermediation, consolidation and capitalisation. *New Political Economy*, *26*(3), 376–388. https://doi.org/10.1080/13563467.2020.1766432

Lopez-Solano, J., & Castañeda, J. D. (2024). 'A promising playground': IDEMIA and the digital ID infrastructuring in Columbia. *Information, Communication, & Society*, *27*(15), 1–17. https://doi.org/10.1080/1369118X.2024.2302995

Luque-Ayala, A., & Neves Maia, F. (2019). Digital territories: Google maps as a political technique in the re-making of urban informality. *Environment and Planning D: Society and Space*, *37*(3), 449–467. https://doi.org/10.1177/0263775818766069

Mahase, E. (2019). Government hands Amazon free access to NHS information. *BMJ*, *367*(1), l6901. https://doi.org/10.1136/bmj.l6901

Mattern, S. (2018, November). Maintenance and care. *Places*. https://placesjournal.org/article/maintenance-and-care/.

McGarry, D. B. (2022, September 22). *Panelists at broadband breakfast event urge the FCC mapping fabric be made public*. Broadband Breakfast. https://broadbandbreakfast.com/panelists-at-broadband-breakfast-event-urge-the-fcc-mapping-fabric-be-made-public/.

Murakami Wood, D., & Mackinnon, D. (2019). Partial platforms and oligoptic surveillance in the smart city. *Surveillance & Society*, *17*(1/2), 176–182. https://doi.org/10.24908/ss.v17i1/2.13116

Murakami Wood, D., & Steeves, V. (2021). Smart surveillance. *Surveillance & Society*, *19*(2), 150–153. https://doi.org/10.24908/ss.v19i2.14916

Narayan, D. (2022). Platform capitalism and cloud infrastructure: Theorizing a hyper-scalable computing regime. *Environment and Planning A: Economy and Space*, *54*(5), 911–929. https://doi.org/10.1177/0308518X221094028

Pasquale, F. (2016). *The black box society: The secret algorithms that control money and information*. Harvard University Press.

Reutter, L., & Åm, H. (2024). Constructing the data economy: Tracing expectations of value creation in policy documents. *Critical Policy Studies*, *18*(4), 1–21. https://doi.org/10.1080/19460171.2023.2300436

Russell, A. L., & Vinsel, L. (2018). After innovation, turn to maintenance. *Technology and Culture*, *59*(1), 1–25. https://doi.org/10.1353/tech.2018.0004

Sadowski, J. (2020). *Too smart: How digital capitalism Is extracting data, controlling our lives, and taking over the world*. MIT Press.

Seaver, N. (2019). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, *24*(4), 421–436. https://doi.org/10.1177/1359183518820366

Sherman, S. (2023). The polyopticon: A diagram for urban artificial intelligences. *AI & SOCIETY*, *38*(3), 1209–1222. https://doi.org/10.1007/s00146-022-01501-3

Taylor, L. (2021). Public actors without public values: Legitimacy, domination and the regulation of the technology sector. *Philosophy & Technology*, *34*(4), 897–922. https://doi.org/10.1007/s13347-020-00441-4

West, S. M. (2019). Data capitalism: Redefining the logics of surveillance and privacy. *Business & Society*, *58*(1), 20–41. https://doi.org/10.1177/0007650317718185

Whittaker, M. (2021). The steep cost of capture. *Interactions*, *28*(6), 50–55. https://doi.org/10.1145/3488666

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.